

12-2015

Population and Ecosystem Attribute Trends of Aquatic Macroinvertebrates

Alyssabeth Beadle
Grand Valley State University

Follow this and additional works at: <https://scholarworks.gvsu.edu/honorsprojects>



Part of the [Life Sciences Commons](#)

ScholarWorks Citation

Beadle, Alyssabeth, "Population and Ecosystem Attribute Trends of Aquatic Macroinvertebrates" (2015).
Honors Projects. 577.
<https://scholarworks.gvsu.edu/honorsprojects/577>

This Open Access is brought to you for free and open access by the Undergraduate Research and Creative Practice at ScholarWorks@GVSU. It has been accepted for inclusion in Honors Projects by an authorized administrator of ScholarWorks@GVSU. For more information, please contact scholarworks@gvsu.edu.

Population and Ecosystem Attribute Trends of Aquatic Macroinvertebrates

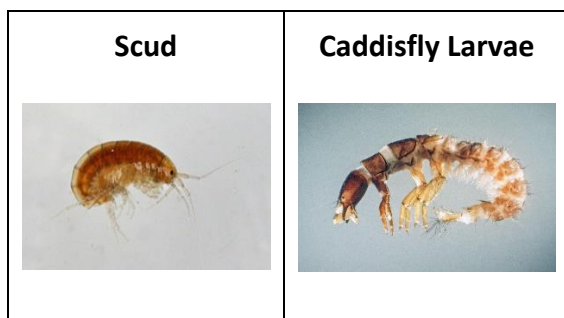
Alyssabeth Beadle

Faculty Mentors: Keith Piccard & John Gabrosek, Ph.D.

Introduction

For the past ten years Professor Keith Piccard of the Grand Valley Biology department has been collecting data on aquatic macroinvertebrates in Allendale Middle School's Sevey drain. He has been studying how precipitation effects aquatic macroinvertebrate population numbers, diversity, and ecosystem attributes. Aquatic macroinvertebrates are defined as organisms without backbones that live on, under, or around rocks and sediment on the bottom of lakes, rivers, and streams. Aquatic macroinvertebrates must be able to be observed by the human eye without the aid of a microscope. The two most common aquatic macroinvertebrates are Scuds and Caddisfly Larvae which are displayed below in Figure One. Scuds have an average length of $\frac{1}{4}$ inch and Caddisfly Larvae generally measure anywhere from $\frac{1}{4}$ inch to $\frac{3}{4}$ inch. Any organisms that did not meet the definition of aquatic macroinvertebrates that were collected from the Sevey drain were not included in the analysis of this study.

Figure One. The Two Most Common Aquatic Macroinvertebrates from our Datasets



The analysis completed for this study will be used by Professor Piccard to further his research which is annually reported to the University of Michigan. This analysis can be used to create an understanding of how precipitation affects population trends, diversity, and ecosystem attributes in the state of Michigan. From his previous research and analysis Professor Piccard hypothesizes that there is a relationship between precipitation and

population numbers, Shannon's Diversity Index, and the ecosystem attributes of the Sevey drain.

Data Collection

The data collection was focused around the various macroinvertebrate species located in the Allendale Public Schools' Sevey Drain. The data was collected by Professor Piccard's sixth grade students at Allendale Middle School. Each student was instructed to record the type of macroinvertebrate observed and the corresponding count for each individual species.

Before I could do any analyses it was necessary to enter and reformat the data so that it could be read into SAS Enterprise Guide. The data entry involved taking the data collection worksheets for each instance of data collection and combining the species counts for yearly population numbers as well as individual specie population numbers. Next, I needed to compile yearly precipitation for Allendale Charter Township. To do this I combined the monthly precipitation amounts for the months July to June for each year to follow the school years that the population counts followed.

The data may have some errors due to inefficiencies in the data collection process and data entry. The collection by sixth grade students creates a possibility for inaccuracies in the data. For example one student in 2006 spotted scuds and listed "lots" as the quantity. This lack of a quantitative value for the specie count caused me to remove these observations from the data. This has an effect on the analyses. Since the data was entered from hard copies into excel there is the possibility for human error in the reporting of the number of each species, either through mistypes or through accidentally omitting worksheets. The data was collected for the

years 2004 to 2015. From year to year there was variation in the types of macroinvertebrates, the overall counts for each species, the diversity level, and the ecosystem attributes.

Methods

As a statistical consultant my role was to provide Professor Piccard with numerical and graphical analyses to see if there truly was any relationship between species counts and precipitation. The first equation was used to predict the overall population number based on the amount of precipitation (in inches) in a year. The second equation used precipitation as the predictor for Shannon's Diversity Index. The next five equations analyzed various ecosystem attributes based on yearly precipitation.

The key variables that were analyzed during the study include yearly precipitation, population reports for each species, and Shannon's Diversity Index, as well as the ecosystem attributes Autotrophy to Heterotrophy Index, Coarse Particulate Organic Matter to Fine Particulate Organic Matter Index, Fine Particulate Organic Matter in Transport to Fine Particulate Organic Matter in Storage in Sediments, Substrate Stability, and Top-down Predator Control. Figure 2 describes each variable in the data set.

Figure Two. Variables & Definitions

Variable	Description
Precipitation	The yearly total rainfall and snowfall combined
Diversity	The Shannon Index for each year
Overall Population Count	The total number of <u>all</u> macroinvertebrates for an entire year
Autotrophy to Heterotrophy Index	The ratio between autotrophic systems and heterotrophic systems
Coarse Particulate Organic Matter to Fine Particulate Organic Matter Index	The ratio of species that eat organic food particles greater than 1.0 mm in size and species that eat organic matter ranging from 0.5 um to 1.0 mm in size
Fine Particulate Organic Matter in Transport to Fine Particulate Organic Matter in Storage in Sediments	The rate at which fine organic matter in the water moves depending on the speed of the water
Substrate Stability	The ratio of species who are scrapers and filtering collectors over species who are shredders and gathering collectors
Top-Down Predator Control	The ratio of the number of species that are predators over all of the other functional feeding groups

The data for yearly precipitation was acquired via the national weather service and is reported in inches. Population numbers for each species were collected by sixth graders at Allendale Public Middle School. This variable is measured by counting each individual found in or around the aquatic environment of the Sevey Drain. To measure diversity in the Sevey Drain, Shannon's Diversity Index, as shown in Figure Three, was calculated for each year. The formula takes into account the number of species recorded so an increase in the index means an

increase in diversity. The values for Shannon's Index range from 0 (no diversity) to 3.5 (highest level of diversity).

Figure Three. Shannon's Diversity Index

$\sum_{i=1}^n (-p_i * \ln(p_i))$	<i>where:</i> <i>p_i = proportion of each species</i> <i>n = total number specie types</i>
----------------------------------	--

The Autotrophy to Heterotrophy Index, as shown in Figure Four, measures the ratio between an autotrophic system, which consists of primarily species that produce autochthonous organic matter as the main energy source (species whose feeding group is scrapers), and a heterotrophic system, which is a system that relies on the input of organic matter as the main energy component (species whose feeding groups are shredders, filtering collectors, and gathering collectors). If the ratio is greater than 0.75 the ecosystem is autotrophic; if the ratio is greater than or equal to 0.75 the ecosystem is heterotrophic.

Figure Four. Autotrophy to Heterotrophy Index

$\frac{\text{Number of Scrapers}}{\text{Number of Shredders} + \text{Number of Collectors}}$
--

Coarse Particulate Organic Matter to Fine Particulate Organic Matter Index measures the ratio of species that eat organic food particles greater than 1.0 mm in size (species whose feeding groups are shredders), such as leaf litter, against species that eat organic matter ranging from 0.5 um to 1.0 mm in size such as organic material that has been broken down (species whose feeding groups are gathering collectors and filtering collectors). Normal values for the Coarse Particulate Organic Matter to Fine Particulate Organic Matter Index, which is

shown in Figure Five, are greater than 0.5 in the fall and winter, and greater than 0.25 in the spring and summer.

Figure Five. Coarse Particulate Organic Matter to Fine Particulate Organic Matter Index

$$\frac{\text{Number of Shredders}}{\text{Number of Filtering + Gathering Collectors}}$$

Fine Particulate Organic Matter in Transport to Fine Particulate Organic Matter in Storage in Sediments is a ratio that determines how the fine organic matter in the water moves depending on the speed that the water is moving. The equation is shown in Figure Six and is a ratio of species that are filtering collectors over species that are gathering collectors. When the ratio is greater than 0.5 the fine particulate organic matter transport is greater than the normal particulate loading in suspension, and when the ratio is less than or equal to 0.5 the normal particulate loading is in suspension.

Figure Six. Fine Particulate Organic Matter in Transport to Fine Particulate Organic Matter in Storage in Sediments

$$\frac{\text{Number of Filtering Collectors}}{\text{Number of Gathering Collectors}}$$

Substrate Stability, whose equation is displayed in Figure Seven, looks at how stable the stream bed and channel are, and how likely they are going to erode away. Substrate Stability is calculated using a ratio of species who are scrapers and filtering collectors over species who are shredders and gathering collectors. When stable substrate is plentiful the ratio is greater than 0.5 and when unstable substrate is plentiful the ratio is less than or equal to 0.5.

Figure Seven. Substrate Stability

$$\frac{\text{Number of Scrapers} + \text{Number of Filtering Collectors}}{\text{Number of Gathering Collectors} + \text{Number of Shredders}}$$

Top-Down Predator Control, whose equation is shown in Figure Eight, indicates a balance between prey species with long and short life cycles. Normal predator to prey balances fall between 0.10 and 0.20, ratios any higher indicate rapid turnover in prey species. The ratio is calculated by dividing the number of species who are predators over the numbers of all other functional feeding groups.

Figure Eight. Top-Down Predator Control

$$\frac{\text{Number of Predators}}{\text{Total Number of all other Functional Feeding Groups}}$$

Professor Piccard hopes to find a relationship between the levels of precipitation during any given year, the overall population count, the diversity levels, and the ecosystem attributes. He hypothesizes that the amount of precipitation for each year can help predict the population numbers, diversity levels, and ecosystem attributes.

Hypothesis 1: There is an association between precipitation in a year and population count.

Hypothesis 2: There is an association between precipitation in a year and Shannon's Diversity Index.

Hypothesis 3: There is an association between precipitation in a year and the Autotrophy to Heterotrophy Index.

Hypothesis 4: There is an association between precipitation in a year and the Coarse Particulate Organic Matter to Fine Particulate Organic Matter Index.

Hypothesis 5: There is an association between precipitation in a year and Fine Particulate Organic Matter in Transport to Fine Particulate Matter in Storage in Sediments.

Hypothesis 6: There is an association between precipitation in a year and Substrate Stability.

Hypothesis 7: There is an association between precipitation in a year and Top-Down Predator Control.

In an attempt to see if precipitation can predict the other variables, we ran seven separate simple linear regressions. All of the linear regressions use precipitation as the independent variable.

Results

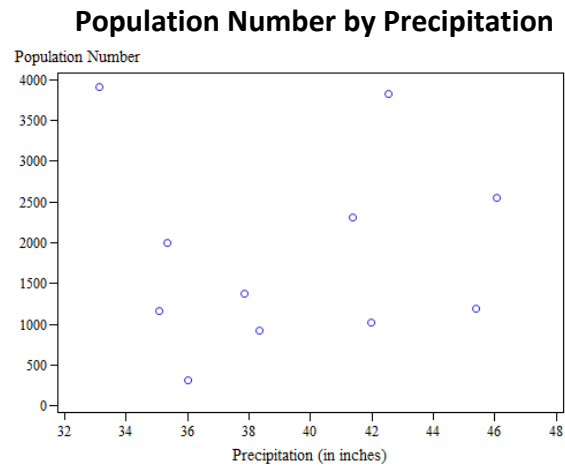
From the sample of macroinvertebrates over an eleven year span of time I was able to analyze the variables precipitation, population counts, diversity, Autotrophy to Heterotrophy Index, Coarse Particulate Organic Matter to Fine Particulate Organic Matter Index, Fine Particulate Organic Matter in Transport to Fine Particulate Organic Matter in Storage in Sediments, Substrate Stability, and Top-down Predator Control. These variables are defined in Figure Three. From this data my plan was to build models that could accurately predict diversity, population numbers, and ecosystem attributes with precipitation as the predictor. To accomplish this I used simple linear regression.

Before I began my analyses I wanted to graph each of my variables against precipitation. These scatter plots are shown in Figure Nine. After looking at each of these scatter plots I saw

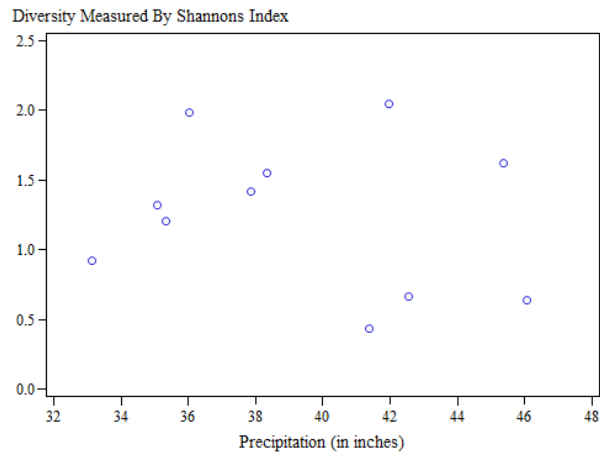
POPULATION AND ECOSYSTEM ATTRIBUTE TRENDS OF AQUATIC MACROINVERTEBRATES

no association between precipitation and the other variables, but to confirm that there was no relationship I went on to run each of the linear regressions.

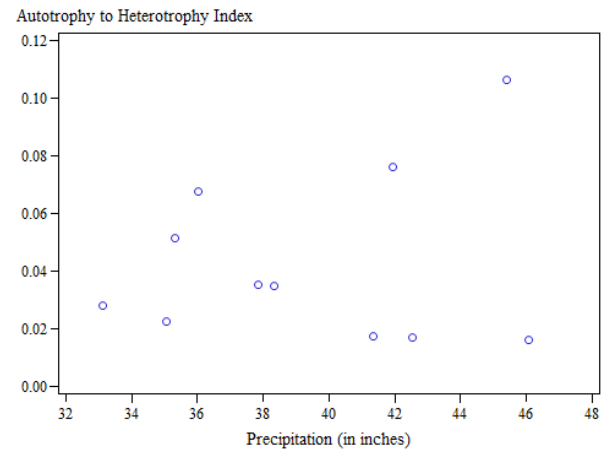
Figure Nine. Scatter Plots with Precipitation as the Independent Variable



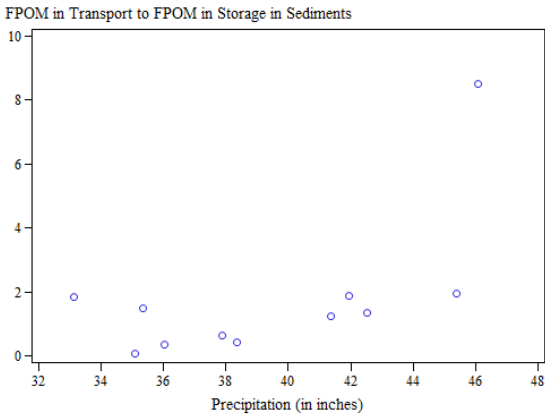
Shannon's Diversity Index by Precipitation



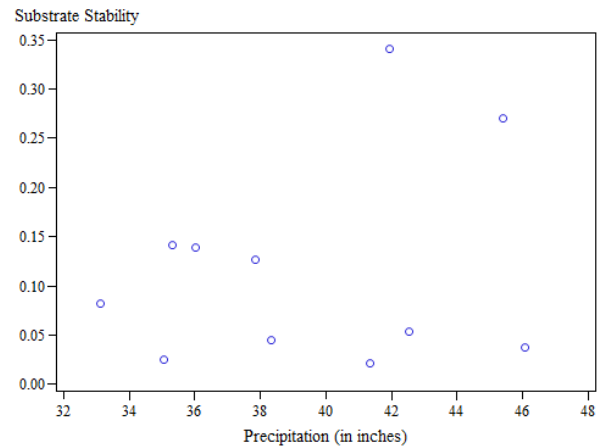
Autotrophy to Heterotrophy Index by Precipitation



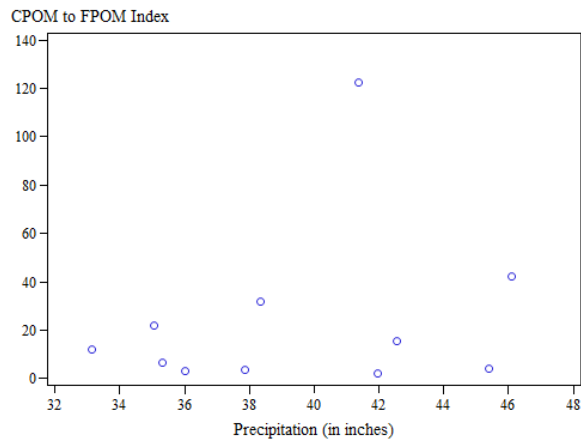
Fine Particulate Organic Matter in Transport to Fine Particulate Organic Matter in Storage in Sediments by Precipitation



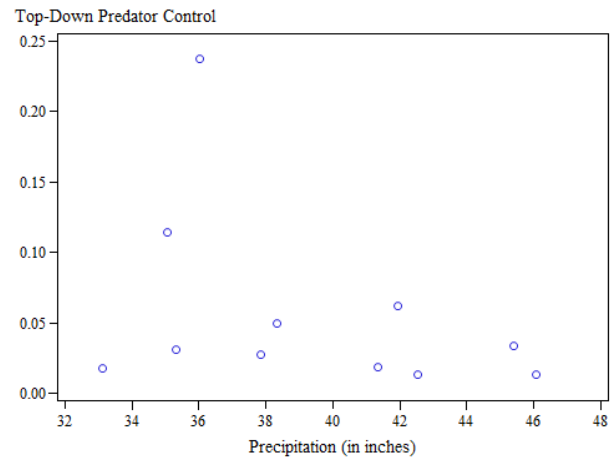
Substrate Stability by Precipitation



Coarse Particulate Organic Matter to Fine Particulate Organic Matter Index by Precipitation



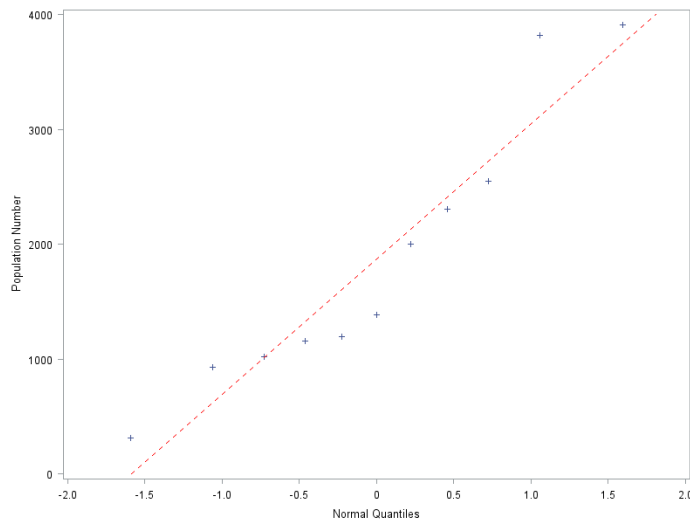
Top-Down Predator Control by Precipitation



The assumptions must be addressed before beginning any analysis on the data. The conditions that must be met for a simple linear regression include the following: a normal population, no extreme errors, and homoscedasticity or equal variances. I am assuming a normal population although our sample size of ten years is not ideal, the points on each of the Q-Q plots lie close to the lines so it is safe to assume normality for all eight of the variables. There were no extreme errors in the dataset and homoscedasticity was assumed. The Q-Q plot

for overall population count is in Figure Ten, the Q-Q plots for diversity, the five ecosystem attributes and precipitation are shown in the appendix. All eight of the plots confirmed that the conditions listed above were met.

Figure Ten. Q-Q Plot for Overall Population Count



The first regression that I ran used population number as the dependent variable and precipitation as the independent variable. The final model for this regression was:

$$\hat{y} = 1456.78 + 10.54(\text{precipitation})$$

with an R^2 value of 0.002. This equation means that as precipitation increases by one inch, population numbers increase by 10.54 macroinvertebrates. The R^2 value means that precipitation accounts for 0.20% of the overall population count variability. It is also important to note that the p-value for this model is 0.9093, which means that there is insufficient evidence to conclude that there is a relationship between overall population count and precipitation. The corresponding output from SAS is located in the appendix.

The second regression used diversity as the dependent variable and precipitation as the independent variable. The final model for this regression was:

$\hat{y}=2.16 - 0.02(\text{precipitation})$ with an R^2 value of 0.034. The R^2 value means that precipitation accounts for 3.4% of the overall population count variability. The p-value for this model is 0.5855, meaning that there is insufficient evidence to conclude that there is a relationship between diversity and precipitation. The corresponding output from SAS is located in the appendix.

The next regression used the Autotrophy to Heterotrophy Index as the dependent variable and precipitation as the independent variable. The regression equation for this model is: $\hat{y} = -0.0064 + 0.0013(\text{precipitation})$ with an R^2 value of 0.035. The R^2 values means that precipitation accounts for 3.5% of the variability in the autotrophy to heterotrophy index. The p-value for this model is 0.5825, indicating that there is insufficient evidence to suggest that there is a relationship between precipitation and the autotrophy to heterotrophy index. The necessary SAS output was included in the appendix.

The regression model that used Fine Particulate Organic Matter in Transport to Fine Particulate Organic Matter in Storage in Sediments as the dependent variable and precipitation as the independent variable had an equation of: $\hat{y} = -10.597 + 0.315(\text{precipitation})$ and has an R^2 value of 0.35. The equation indicates that for every one inch increase in precipitation the Fine Particulate Organic Matter in Transport to Fine Particulate Organic Matter in Storage in Sediments increases by 0.315. The R^2 value of 0.35 indicates that precipitation accounts for only 35% of all of the variability in Fine Particulate Organic Matter in Transport to Fine Particulate Organic Matter in Storage in Sediments. The model has a p-value of 0.054 so at the significance level of 0.05 there is insufficient evidence to suggest that there is a relationship between

precipitation and fine particulate organic matter in transport to fine particulate matter in storage in sediments. The relevant SAS output is included in the appendix.

The next regression that I ran used Coarse Particulate Organic Matter to Fine Particulate Organic Matter Index as the dependent variable with precipitation as the independent variable. This regression had an equation of: $\hat{y} = -50.501 + 1.8996(\text{precipitation})$ with an R^2 of .056. The R^2 value means that precipitation explains only 5.6% of all of the variability in Coarse Particulate Organic Matter to Fine Particulate Organic Matter Index. The p-value for this model is .453 which means that there is insufficient evidence to conclude that there is a relationship between Coarse Particulate Organic Matter to Fine Particulate Organic Matter Index and precipitation. All relevant SAS output is provided in the appendix.

To determine if there was a relationship between substrate stability and precipitation I ran a regression using substrate stability as the dependent variable and precipitation as the independent variable. The model for the regression is as follows:

$\hat{y} = -0.106 + 0.0057(\text{precipitation})$ with an R^2 value of 0.056. The R^2 value means that precipitation explains only 5.6% of all variability in substrate stability. It is important to note that the p-value for this model is 0.48 which means that there is insufficient evidence to suggest that there is a relationship between precipitation and substrate stability. All necessary SAS output is included in the appendix.

The final regression that I ran used top-down predator control as the dependent variable and precipitation as the independent variable. The final model for this regression was: $\hat{y} = 0.281 - 0.0057(\text{precipitation})$ with an R^2 value of 0.138. The R^2 value indicates that precipitation explains 13.8% of all variability in top-down predator control. It is important to

note that with a p-value of 0.2598 there is insufficient evidence to suggest that there is a relationship between precipitation and top-down predator control. The corresponding SAS output is located in the appendix.

All of these models are not significant, which is to say that none of them show sufficient evidence that precipitation has a relationship with any of the variables. This is important to remember when referencing these equations, since none of these equations accurately make predictions using precipitation.

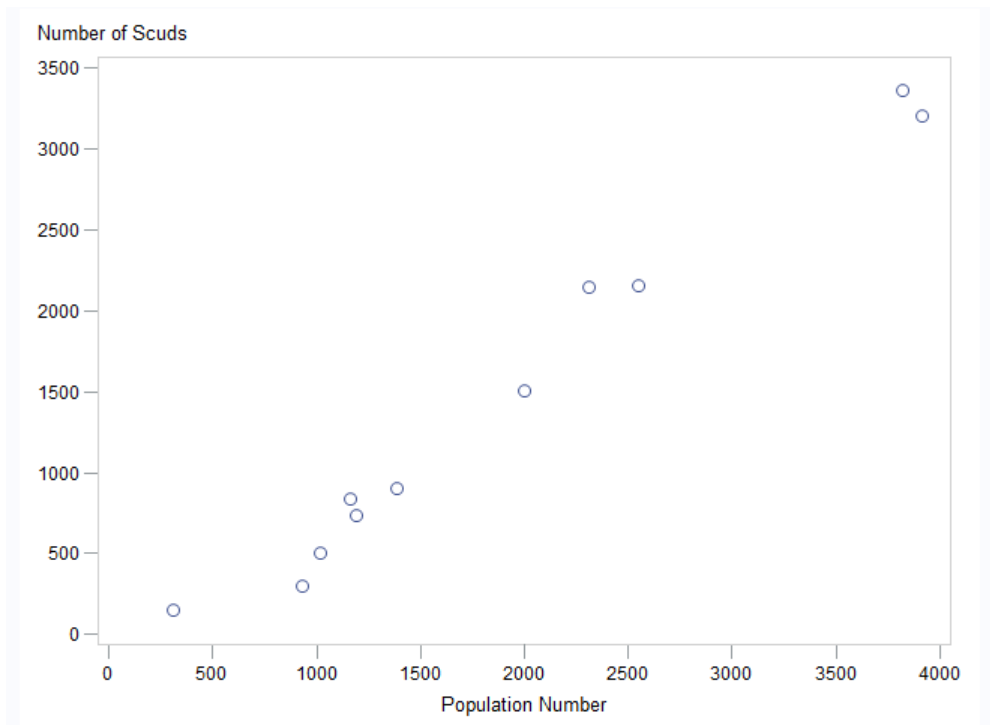
Figure Eleven. Regression Results

Dependent Variable	Independent Variable	R²	Regression Equation
Population Number	Precipitation	0.002	$\hat{y} = 1456.78 + 10.54(\text{precipitation})$
Shannon's Diversity Index	Precipitation	0.034	$\hat{y} = 2.16 - 0.02(\text{precipitation})$
Autotrophy to Heterotrophy Index	Precipitation	0.035	$\hat{y} = -0.0064 + 0.0013(\text{precipitation})$
Fine Particulate Organic Matter in Transport to Fine Particulate Organic Matter in Storage in Sediments	Precipitation	0.35	$\hat{y} = -10.597 + 0.315(\text{precipitation})$
Coarse Particulate Organic Matter to Fine Particulate Organic Matter Index	Precipitation	0.056	$\hat{y} = -50.501 + 1.8996(\text{precipitation})$
Substrate Stability	Precipitation	0.056	$\hat{y} = -0.106 + 0.0057(\text{precipitation})$
Top-Down Predator Control	Precipitation	0.138	$\hat{y} = 0.281 - 0.0057(\text{precipitation})$

While doing the data entry I noticed that there was a high proportion of the number of scuds out of the total population count every year. I was wondering whether as the number of scuds increased the population count also increased. Because the number of scuds contributed

to the total population count I could not complete a regression but I did choose to graph the number of scuds by the population count as shown in Figure Twelve. The graph shows that scuds do in fact have a large impact on the population count.

Figure Twelve. Scuds by Population Count



Discussions/Recommendations

It was hypothesized by Professor Piccard that there was a relationship between precipitation and population count, diversity, and ecosystem attributes. During my analysis I found no significant evidence to prove that there is any relationship between those variables. This causes me to conclude that there is insufficient evidence that precipitation can predict population numbers, Shannon's Diversity Index, Autotrophy to Heterotrophy Index, Coarse Particulate Organic Matter to Fine Particulate Organic Matter Index, Fine Particulate Organic

Matter in Transport to Fine Particulate Organic Matter in Storage in Sediments, Substrate Stability, and Top-down Predator Control.

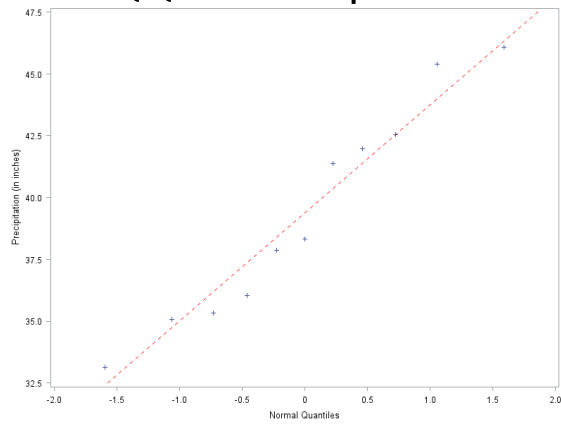
While doing this analysis I encountered a few limitations. Because Professor Piccard has only been gathering the data for 11 years I was unable to look further into the past to add more data to create a larger data set that would be more reliable and less influenced by outlier years. If one chooses to repeat this study, I would recommend that they should include more years to better show if there is a relationship between the variables. I would also recommend that the data is collected and recorded in way that will create fewer errors.

Acknowledgments

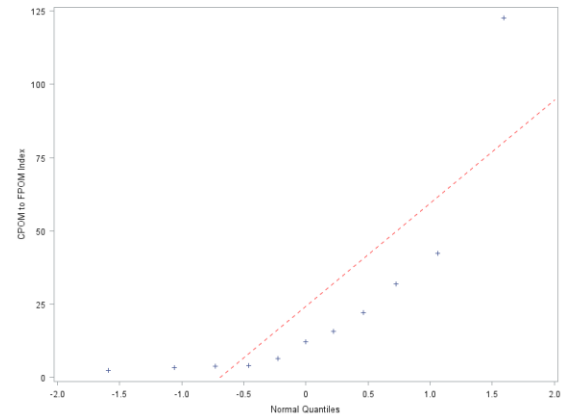
I would like to thank Mackenzie Kohler for her work on the STA 319 project with me, without her help I would not have been able to complete this project. I would also like to thank Professor John Gabrosek for teaching me how to be an effective statistical consultant and providing guidance to me on this project. A special thanks to Professor Keith Piccard for making the data available to me and teaching me the biology behind my analyses.

Appendix

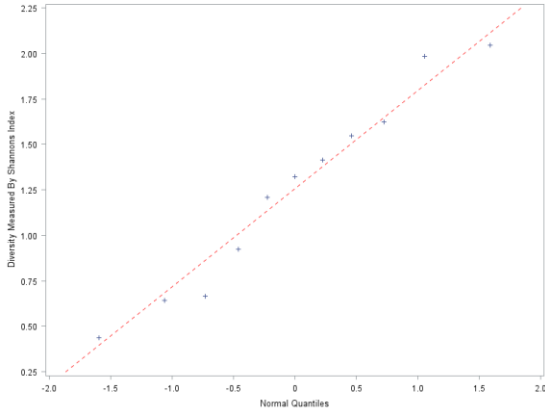
Q-Q Plot for Precipitation



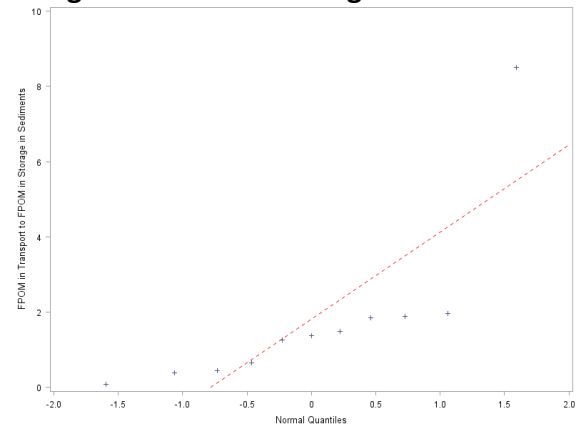
Q-Q Plot for Coarse Particulate Organic Matter to Fine Particulate Organic Matter Index



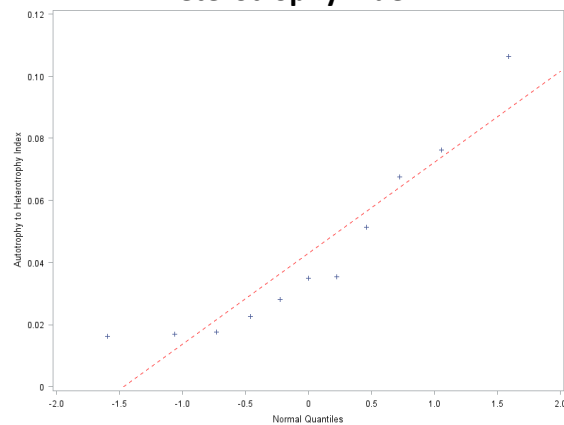
Q-Q Plot for Shannon's Index



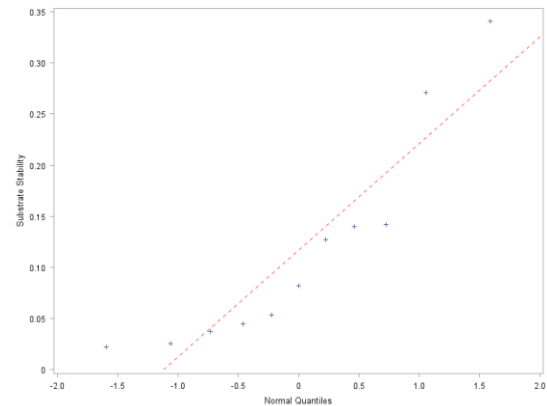
Q-Q Plot for Fine Particulate Organic Matter in Transport to Fine Particulate Organic Matter in Storage in Sediments

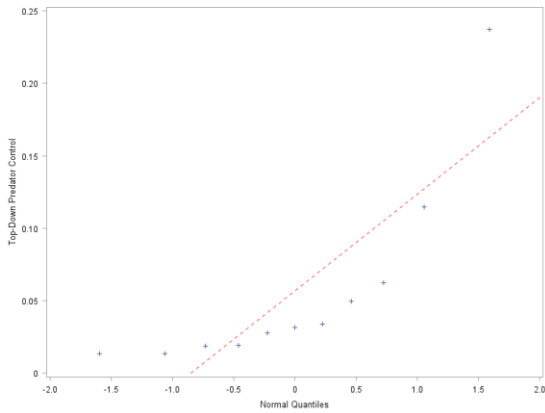


Q-Q Plot for Autotrophy to Heterotrophy Index



Q-Q Plot for Substrate Stability



Q-Q Plot for Top-Down Predator Control**Regression Output for Population Count as Dependent Variable**

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	21123.14	21123.14	0.01	0.9093
Error	9	13838373.41	1537597.05		
Corrected Total	10	13859496.55			

R-Square	Coeff Var	Root MSE	population Mean
0.001524	66.25212	1239.999	1871.636

Parameter	Estimate	Standard Error	t Value	Pr > t
Intercept	1456.783128	3559.147734	0.41	0.6919
precipitation	10.537050	89.900294	0.12	0.9093

Regression Output for Diversity as Dependent Variable

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	0.09981119	0.09981119	0.32	0.5855
Error	9	2.80762441	0.31195827		
Corrected Total	10	2.90743559			

R-Square	Coeff Var	Root MSE	diversity Mean
0.034330	44.45246	0.558532	1.256471

Parameter	Estimate	Standard Error	t Value	Pr > t
Intercept	2.158259841	1.60314571	1.35	0.2111
precipitation	-0.022904956	0.04049376	-0.57	0.5855

Regression Output for Autotrophy to Heterotrophy Index as Dependent Variable

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	0.00029946	0.00029946	0.33	0.5825
Error	9	0.00829022	0.00092114		
Corrected Total	10	0.00858968			

R-Square	Coeff Var	Root MSE	AutoHetero1 Mean
0.034863	70.53647	0.030350	0.043028

Parameter	Estimate	Standard Error	t Value	Pr > t
Intercept	-0.0063673155	0.08711371	-0.07	0.9433
precipitation	0.0012546067	0.00220040	0.57	0.5825

Regression Output for Coarse Particulate Organic Matter to Fine Particulate Organic Matter Index as Dependent Variable

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	686.48714	686.48714	0.53	0.4850
Error	9	11650.18669	1294.46519		
Corrected Total	10	12336.67382			

R-Square	Coeff Var	Root MSE	OrganicMatter2 Mean
0.055646	148.1393	35.97868	24.28706

Parameter	Estimate	Standard Error	t Value	Pr > t
Intercept	-50.50086533	103.2689913	-0.49	0.6365
precipitation	1.89957326	2.6084651	0.73	0.4850

Regression Output for Fine Particulate Organic Matter in Transport to Fine Particulate Matter in Storage in Sediments as Dependent Variable

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	18.88164554	18.88164554	4.89	0.0543
Error	9	34.75701271	3.86189030		
Corrected Total	10	53.63865826			

R-Square	Coeff Var	Root MSE	FPOM3 Mean
0.352016	108.7822	1.965169	1.806517

Parameter	Estimate	Standard Error	t Value	Pr > t
Intercept	-10.59672483	5.64059235	-1.88	0.0930
precipitation	0.31503570	0.14247538	2.21	0.0543

Regression Output for Substrate Stability as Dependent Variable

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	0.00611816	0.00611816	0.53	0.4832
Error	9	0.10294682	0.01143854		
Corrected Total	10	0.10906498			

R-Square	Coeff Var	Root MSE	Substrate4 Mean
0.056096	91.55126	0.106951	0.116821

Parameter	Estimate	Standard Error	t Value	Pr > t
Intercept	-.1064466509	0.30697991	-0.35	0.7367
precipitation	0.0056708783	0.00775399	0.73	0.4832

Regression Output for Top-Down Predator Control as Dependent Variable

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	0.00617901	0.00617901	1.45	0.2598
Error	9	0.03845225	0.00427247		
Corrected Total	10	0.04463126			

R-Square	Coeff Var	Root MSE	TopDown5 Mean
0.138446	115.8489	0.065364	0.056422

Parameter	Estimate	Standard Error	t Value	Pr > t
Intercept	0.2807970375	0.18761362	1.50	0.1687
precipitation	-.0056990084	0.00473892	-1.20	0.2598

References

"Caddisfly Larvae." Web. 7 Dec. 2015. <<http://www.state.ky.us/nrepc/water/cadfly.htm>>.

"Scud." *Scud*. Web. 7 Dec. 2015. <<http://www.fcps.edu/islandcreekes/ecology/scud.htm>>.

"Temperature - Precipitation - Sunshine - Snowfall." *Climate Grand Rapids*. Web. 12 Nov. 2015.

<<http://www.usclimatedata.com/climate/grand-rapids/michigan/united-states/usmi0344/2011/12>>.